

# Super-resolution wavefront reconstruction

Sylvain Oberti<sup>1</sup>, Carlos Correia<sup>2</sup>, Thierry Fusco<sup>3,4</sup>, Benoit Neichel<sup>3</sup>, and Pierre Guiraud<sup>5</sup>

<sup>1</sup> European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching bei Muenchen, Germany  
e-mail: soberti@eso.org

<sup>2</sup> Space ODT – Optical Deblurring Technologies Unip Lda, Porto, Portugal

<sup>3</sup> Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France

<sup>4</sup> DOTA, ONERA, Université Paris Saclay, 91123 Palaiseau, France

<sup>5</sup> Instituto de Ingeniería Matemática and Centro de Investigación y Modelamiento de Fenómenos Aleatorios – Valparaíso, Universidad de Valparaíso, Valparaíso, Chile

Received 4 May 2022 / Accepted 2 August 2022

## ABSTRACT

**Context.** Cutting-edge, ground-based astronomical instruments are fed by adaptive optics (AO) systems that are aimed at providing high performance down to the visible wavelength domain on 10 m class telescopes and in the near infrared for the first generation instruments of Extremely Large Telescopes (ELTs). Both applications lead to a large ratio between the telescope diameter,  $D$ , and the coherence length or Fried parameter,  $r_0$ , that is  $D/r_0$ . As the parameter that defines the required number of degrees of freedom of the AO system,  $D/r_0$  drives the requirement to reconstruct the incoming wavefront with ever-higher spatial resolution. In this context, super-resolution (SR) appears as a potential game changer. Indeed, SR promises to dramatically expand the range of spatial frequencies that can be reconstructed from a set of lower resolution measurements of the wavefront.

**Aims.** As a technique that seeks to upscale the resolution of a set of measured signals, SR retrieves higher-frequency signal content by combining multiple lower resolution sampled data sets. It is well known both in the temporal and spatial domains and widely used in imaging to reduce aliasing and enhance the resolution of coarsely sampled images. This study applies the SR technique to the bi-dimensional wavefront reconstruction. In particular, we show how SR is intrinsically suited for tomographic multi-wavefront sensor (WFS) AO systems, revealing many of its advantages with minimal design effort.

**Methods.** We provide a direct space and Fourier optics description of the wavefront sensing operation and we demonstrate how SR can be exploited through signal reconstruction, especially within the framework of periodic non-uniform sampling. We investigate both meta-uniform and non-uniform sampling schemes and we show that under some conditions, both sampling schemes enable a perfect reconstruction of band-limited signals. We also provide a SR bi-dimensional model for a Shack-Hartmann (SH) WFS, along with an analysis of the characteristics of the sensitivity function. We validated the SR concept with numerical simulations of representative multi-WFS SH AO systems. Finally, we explored the extension of the method to pyramid WFSs.

**Results.** Our results show that combining several WFS samples in a SR framework grants access to a greater number of modes than the native one offered by a single WFS (despite the fixed sub-aperture size across samples). We show that the wavefront reconstruction achieved with four WFSs can be equivalent to a single WFS providing a sampling resolution that is twice greater (linear across the telescope aperture). We also show that the associated noise propagation is not degraded under SR. Finally, we show that the concept can be extended to the signal produced by single pyramid WFS, with its four re-imaged pupils serving as multiple non-redundant samples.

**Conclusions.** We find that SR applied to wavefront sensing and reconstruction (WFR) offers a new parameter space to explore, as it decouples the size of the sub-aperture from the desired wavefront sampling resolution. By shifting away from outdated assumptions, new and more flexible, better-performing AO designs have now become possible.

**Key words.** techniques: high angular resolution – instrumentation: adaptive optics – instrumentation: high angular resolution – turbulence – atmospheric effects – techniques: image processing

## 1. Introduction

The term ‘super resolution’ (SR), as applied in this work, is a technique found to serve the needs of the computational imaging field, whereby the joint application of optical design and signal processing techniques is engineered to obtain higher-resolution (HR) data products from multiple low-resolution (LR) samples. Such SR techniques are relatively standard in vision science whenever image data is available with sub-pixel shifts between LR samples. The first work on image upscaling was published in 1984 (Tsai & Huang 1984) and the term ‘super resolution’ itself appeared around 1990 (Irani & Peleg 1990). Park et al. (2003) provides an excellent review, a sign of how, two decades on, SR imaging has been established as a prominent technique.

Such SR techniques are not new in the astronomical observation and instrumentation communities. One relevant example is the ‘drizzle’ method developed to enhance the resolution of under-sampled HST images from Fruchter & Hook (2002), where the high spatial frequency information of the HST images are recovered by combining sub-pixel dithered images. In this paper, we explore the application of SR to wavefront reconstruction for astronomical adaptive optics (AO), as proposed in the presentation from Oberti et al. (2017).

Reconstructing wavefronts instead of focal plane images is particularly interesting. Indeed, in the field of AO assisted astronomy instruments, the optical systems are typically designed to achieve a close to Nyquist–Shannon sampling of the diffraction limited image. In other words, the images are well sampled and

the potential benefits from SR are moderate. On the other hand, the pupil image is typically not limited by diffraction in current AO systems. Specifically, the field stop is sufficiently opened to allow for the set of spatial frequencies of interest to be sensed by the WFS. Moreover, the bi-dimensional wavefront surface can be described by a theoretically unlimited set of spatial frequencies, hence, pushing the SR limit down to extremely small scales.

It turns out that the reconstruction problem is often ill-conditioned because of insufficient LR data or is based on an ill-posed system model. Consequently, some degree of regularisation is required to make it solvable, providing realistic and stable solutions. Even in the absence of multiple LR samples, the use of regularisation provides a means to estimate signal content beyond the native Nyquist–Shannon frequency. We call this strategy ‘statistical SR’, which has by now become well-established in AO. Indeed, the use of regularised reconstruction in the form of a minimum mean square error (MMSE) or maximum a posteriori (MAP) stochastic estimator (Fusco et al. 2001; Conan 2014; Correia & Teixeira 2014); additionally and more generally, methods that draw upon Tikhonov regularisation are now commonplace. Furthermore, SR can also be approached ‘geometrically’, which we aim to explore in this work. This approach is based on the idea that the samples from multiple wavefront measurements can be combined to allow for the reconstruction of higher-resolution spatial frequencies than what could be achieved by each single sample alone. Multiple sampled wavefront measurements (mostly across space, however, we could similarly use video sequences across time or a mix of both) are thus combined from all available lower-resolution WFS measurements, provided that each contains some form of unique phase information. In practice, however, each SR facet can only be interpreted asymptotically: it is ‘geometric’ when we approach the limit of least-squares reconstruction and ‘statistical’ when fewer and fewer samples are used.

At this stage, the following observation is paramount: multiple frames are natively granted for SR in classically designed tomographic systems, where multiple WFSs are used, each looking at a single guide-star on a distinctive line of sight. The multitude of diverse WFSs back-projections at different ranges along those lines of sight sample the turbulence in altitude at offset locations (grid points), which is sufficient to perform SR at every altitude range. The exception is the pupil plane where all sampling grids perfectly overlap, making the relative phase information vanish. Depending on the geometric configuration, this may be the case as well at certain altitudes for which the relative grid offsets are integer multiples of one sampling step.

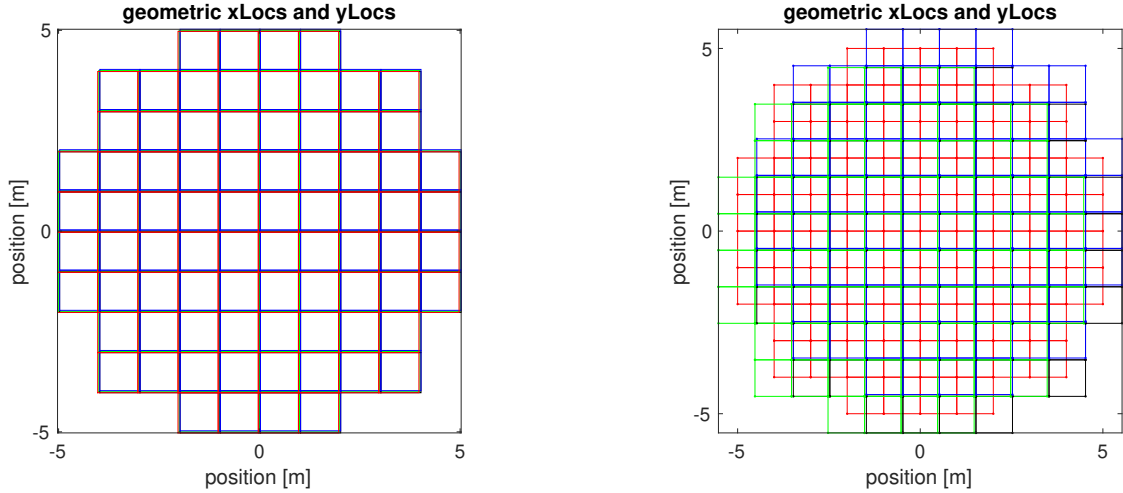
Building on the footsteps of Oberti et al. (2017), the SR concept has been implemented and investigated for several AO applications. For instance, the shape of a high-order DM has been estimated by reconstructing a posteriori a temporal sequence of LR measurements from a single SH WFS, with proper fractional sub-aperture shifts applied to each temporal sample (Willez et al. 2019). Actually, such synthetic aperture wavefront sensing technique had already been applied in the field of eye aberrometry by Bara et al. (2013).

However, the main applications of SR relate to tomographic AO systems where the signal from multiple WFS are combined. A first prominent example include Ellerbroek & Rigaut (2001) and Wang et al. (2012) who tomographically reconstructed the 3D wavefront with an over-sampling factor of up to 2 from many identical WFS conjugated to the pupil-plane. However, the principles of SR (‘statistical SR’ in this case) were not drawn upon at that time.

In Oberti et al. (2017), numerical simulations were used to describe the application of SR to GALACSI NFM (Oberti et al. 2018), the laser tomography AO (LTAO) mode in operation at ESO’s Very Large Telescope (VLT). In fact, SR is already built-in the tomographic reconstruction process as the four WFSs exhibit slight relative mis-registrations that are not compensated for by hardware; instead, these are calibrated and taken into account in the system model, the tomographic interaction matrix. Hence, SR is enabled through the MMSE reconstruction process. More recently, the ELT Multi Conjugate Adaptive Optics (MCAO) system formerly named MAORY now MORFEO, incorporated SR in its design. The concept has been analyzed via end-to-end simulations, with the first studies enabling SR in altitude layers without adding misregistrations in the pupil plane, such as in the works by Oberti et al. (2019) and Busoni et al. (2019). Another study taking advantage of the full ‘geometrical’ SR by introducing voluntary shifts and rotations between the sampling grids was carried out by Agapito et al. (2020). With such misregistrations (characterised by a diversity of sampling geometries), super-resolved wavefronts can be reconstructed in all layers, including the DM plane. As a result, it becomes possible to choose to either gain in aliasing reduction at the iso number of sub-apertures or to reduce the number of sub-apertures, while preserving a similar level of performance. In this case, SR opens the door to new possibilities for WFS design choices that allow for an increase in the sub-aperture size and fitting more pixels per sub-aperture, for example, when dealing with LGS spot elongation. Consistent results have been obtained for several other tomographic AO systems, as, for instance, the HARMONI Laser Tomography Adaptive Optics (LTAO) mode (Fusco et al. 2022) the Giant Magellan Telescope (GMT) GLAO and LTAO systems (Van Dam 2017, priv. comm.), the future VLT MCAO system MAVIS (Cranney et al. 2021; Rigaut et al. 2020), and KAPA, the Keck Observatory LTAO system (Wizinowich et al. 2020).

A meta-uniform sampling can be simply achieved with four sampling grids by introducing proper offsets of half a sub-aperture width, as depicted in Fig. 1. It appears clearly that with grids of finite size, the uniformity is not granted at the edge of the aperture. To enable SR at the ground but also over most conjugation altitudes, one obvious way is to break sampling and reconstruction symmetries (or regularities), and diversify the measurements (i.e. avoid redundancy) by adopting a non-uniform sampling. The most straightforward way is to offset and rotate the sampling grids with respect to one another. Other non-uniform sampling options can be envisioned, with some examples presented in Appendix B.

This method is meant to support the control of a  $N_{\text{act}} \times N_{\text{act}}$  degree-of-freedom deformable mirror (DM) based on measurements of several WFS each of a number of sampling elements,  $N_{\text{sa}} \times N_{\text{sa}}$ , much smaller than  $N_{\text{act}}$ , that is, typically  $N_{\text{sa}} \approx \{1/2 - 2/3\}N_{\text{act}}$ , depending on the number of WFS. SR greatly relaxes the requirements on WFS alignment. For our application domain, we break free from registering the DM to the WFS as accurately as possible, yet once we have completed this step, we need to know the registration precisely in order to forward model this information and estimate higher order data. In other words, SR allows us to trade accuracy for precision; thus, while this distinction may seem elusive, it is actually fundamental to the process. We note that the registration must indeed be calibrated precisely – but not more precisely than what classical AO systems require. Indeed, Agapito et al. (2020) suggests that the super-resolved cases are robust to model calibration errors: the performance is less sensitive to an error in the knowledge of



**Fig. 1.** Illustration of four lenslet arrays, aligned to the *left* and shifted by 1/2 of a sub-aperture to the *right*, where the four colours used (red, green, blue, and black) are no longer superimposed.

the introduced misregistration amplitude than that of the classical MCAO system with co-aligned lenslet array grids in the pupil plane. Furthermore, for multi-WFS AO systems, asymmetric sampling may reduce the unseen or badly-seen modes (Neichel et al. 2009).

It does not go unnoticed that there are certain commonalities with other phase diverse techniques in imaging and AO. While classical phase-diversity relies on introducing a longitudinal phase encoding in the wavefront plane along the propagation path (Gerchberg 1974), the technique we explore here introduces it transversely with respect to the propagation. The SR technique operates on fused data from all available lower-resolution WFS measurements provided each contains some form of unique phase information – the basic prerequisite for achieving super resolution. The unique phase information can be used to separate the aliased high-frequency from the content in low-frequencies of interest, and the higher-resolution wavefront can be accurately reconstructed. Multiple samples can be acquired sequentially following the same principle, as long as the observed scene (the wavefront in our case) remains static.

In the following, the SR technique is first described from a theoretical point of view in the general and mono-dimensional case (Sect. 2). In particular, we highlight the provided benefits in terms of sensitivity to higher spatial frequencies and aliasing reduction. We show that the signal can be reconstructed without aliasing with uniform or non-uniform samples (under some conditions) at the Nyquist–Shannon average sampling rate. Then, the bi-dimensional case is investigated further through analytical models (Sect. 3.1). For SH WFSs, the sensitivity functions (Sect. 3.2) are studied in the diffraction limited and extended source cases. In Sect. 3.3, the application of SR to multi WFS AO systems is then analysed via numerical models by evaluating the reconstruction error and the noise propagation. Finally, we propose an extension of the SR concept to a single WFS, in the specific case of the pyramid WFS (Sect. 4).

## 2. Super-resolution principle

In this section, for the sake of clarity, we first recall the super-resolution principle in the case of a mono-dimensional signal sampled by multiple periodic grids. Then, the main results applicable to AO are highlighted, in the direct space and the Fourier domain. This chapter provides a general description in the sense

that the described SR technique can be applied to any continuous signal. In the specific case of adaptive optics, it may be applied to any kind of WFS, measuring the phase, its derivative, Laplacian, or another linear transformation of the phase.

### 2.1. Super-resolution framework in the direct space

#### 2.1.1. Combination of several sampling grids: general case

Let  $u := \{u_n\}_{n \in \mathbb{Z}}$  be a sequence of  $\mathbb{R}$  and  $f : \mathbb{R} \rightarrow \mathbb{R}$  a real function. If  $f$  is discretely sampled at the points of the sequence  $u$ , then the result of the sampling is represented by the distribution:

$$f_u(x) := \sum_{n=-\infty}^{+\infty} f(x) \delta(x - u_n) \quad \forall x \in \mathbb{R}. \quad (1)$$

It follows from Eq. (1) that if we own  $N$  sampling grids  $u^1, u^2, \dots, u^N$ , and that if  $w$  is any sequence whose set of points is the union of the points of  $u^1, u^2, \dots, u^N$ , then

$$f_w(x) = \sum_{\ell=1}^N f_{u^\ell}(x) \quad \forall x \in \mathbb{R}. \quad (2)$$

In other words, the sum of  $N$  sampling of  $f$  at the points of the sequences  $u^1, u^2, \dots, u^N$  is a sampling of  $f$  at the union of the sequences, that is,  $f_w$ . For the AO scientist, each sampling sequence stands for the sampling grid corresponding to one WFS, composed of sampling points in the wavefront space, basically located at the center of each sub-aperture. In the following, without loss of generality and for any considered dimension, we use the term ‘sampling grid’ instead of ‘sampling sequence’. The quantity  $N$  stands for the number of WFS samples at play, namely, the number of WFSs or the number of temporal wavefront measurements that are combined. The sampling cell – which we call a ‘sub-aperture’ – is nominally the squared sub-aperture that averages the wavefront within its bounds.

#### 2.1.2. Periodic non-uniform sampling

Following our brief introduction to the general framework, we go on to consider a case of particular interest for the AO community: the periodic non-uniform sampling (PNS) case. The PNS

consists in combining several uniform sampling grids into a single non-uniform meta-sampling grid. The PNS is a scheme of deterministic sampling that was first introduced by [Kohlenberg \(1953\)](#).

Let us consider  $N$  uniform sampling grids,  $u^1, u^2, \dots, u^N$ , with the same sampling step,  $h$ . In AO, this corresponds to a set of  $N$  WFSs with  $h$  as the spatial distance between the centers of two neighbouring sub-apertures. The sampling step,  $h$ , is classically equal to the sub-aperture width. The first sampling grid  $u^1$  is given by:

$$u_n^1 = nh + u_0^1 \quad \forall n \in \mathbb{Z},$$

where  $u_0^1 \in \mathbb{R}$  is the reference origin, which can be for instance the center of the telescope pupil. For any  $\ell \in \{1, \dots, N\}$ , the sampling grid,  $u^\ell$ , is obtained by applying a translation of length  $\Delta_\ell$  to  $u^1$ . For the AO scientist,  $\Delta_\ell$  is equivalent to a shift of the WFS sampling grid number  $\ell$  relative to an arbitrary reference WFS, whose  $\ell = 1$ . So, we have  $\Delta_1 = 0$  and we suppose  $\Delta_\ell \in (0, h)$  if  $\ell \neq 1$ . Then, for every  $\ell \in \{1, \dots, N\}$ , we have:

$$u_n^\ell = nh + u_0^1 + \Delta_\ell \quad \forall n \in \mathbb{Z}.$$

Let  $w$  be any sequence whose points are the union of those of the sampling grids  $u^1, u^2, \dots, u^N$ . For example,  $w$  can be taken such that

$$u_q^\ell = w_{qN+\ell-1} \quad \forall q \in \mathbb{Z} \quad \text{and} \quad \ell \in \{1, \dots, N\}. \quad (3)$$

In this way, if  $\Delta_1 < \Delta_2 < \dots < \Delta_N$ , the sequence  $\{w_n\}_{n \in \mathbb{Z}}$  is sorted in increasing order (its points are labelled according to their spatial order) and if  $n - n' = N$ , then  $w_n$  and  $w_{n'}$  are two successive points of a same sampling grid,  $u^\ell$ . It is interesting to note here that the operation described in Eq. (2) resembles a spatial reconstruction process, whereby the sampling points are re-arranged with respect to their spatial coordinates. In other words, the sampling sequences are not concatenated but, rather, organized according to the order of the sampling points. A simple toy example is easy to grasp: a sinusoid that aliases on each WFS, when samples from multiple WFSs are rearranged, is no longer seen to alias (depending on  $N$ ,  $\Delta_\ell$  and the actual frequency of the sinusoid).

### 2.1.3. From fractional sample shifts to meta-uniform sampling

In the case where the shifts  $\Delta_\ell$  are multiple of  $\frac{h}{N}$ , that is:

$$\Delta_\ell = (\ell - 1) \frac{h}{N} \quad \forall \ell \in \{1, 2, \dots, N\}, \quad (4)$$

then Eq. (3) can be expressed as:

$$w_{qN+\ell-1} = u_q^\ell = qh + u_0^1 + (\ell - 1) \frac{h}{N} = (qN + \ell - 1) \frac{h}{N} + u_0^1,$$

for every  $q \in \mathbb{Z}$  and  $\ell \in \{1, \dots, N\}$ . Therefore,

$$w_n = n \frac{h}{N} + u_0^1 \quad \forall n \in \mathbb{Z}.$$

Then,

$$\sum_{\ell=0}^{N-1} f_{u^\ell}(x) = f_w(x) = \sum_{n=-\infty}^{+\infty} f(x) \delta\left(x - n \frac{h}{N} - u_0^1\right) = (f \cdot \text{III}_{\frac{h}{N}})(x), \quad (5)$$

where

$$\text{III}_{\frac{h}{N}}(x) := \sum_{n=-\infty}^{+\infty} \delta\left(x - \frac{nh}{N} - u_0^1\right)$$

is the Dirac comb of period  $h/N$  centered at  $u_0^1$ .

Let us now consider a superposition of  $N$  sampling grids composed of sub-apertures of a size,  $h$ . In each sub-aperture, the function  $f = \bar{s}_h$  is the result of averaging the signal,  $s$ , by means of a convolution with the top hat function  $\chi_h : \mathbb{R} \rightarrow \mathbb{R}$ , which is the uniform density over the interval  $I_h := [-h/2, h/2]$ ,

According to Eq. (5), the signal sampled by the superposition of sampling grids, that we call ‘meta-uniform sampling’, is thus given by the distribution:

$$s_{N,h}(x) = (\bar{s}_h \cdot \text{III}_{\frac{h}{N}})(x) = (s * \chi_h) \cdot \text{III}_{\frac{h}{N}}(x) \quad \forall x \in \mathbb{R}. \quad (6)$$

We note that Eq. (6) suggests that we can decouple the size of the sub-aperture,  $h$ , from the effective sampling step,  $h/N$ , that is, that we can generate a ‘meta-sampler’ with regular sampling points located every  $h/N$  and whose sub-apertures have a size that is different from the sampling step; this differs from classical AO designs, where the spatial sampling step and the sub-aperture width are the same.

### 2.2. Mono-dimensional representation in the Fourier domain

The Fourier transform of Eq. (6) is expressed as:

$$\hat{s}_{N,h}(k) = h \hat{s}(k) \text{sinc}(\pi h k) * \text{III}_{\frac{h}{N}}(k), \quad (7)$$

with  $k$  as the spatial frequency in  $\text{m}^{-1}$  when considering the wavefront sensing-related spatial sampling case. The Fourier transform of the function  $s$  is multiplied by the sinc function that depends only on the sub-aperture size,  $h$ . The sinc function can be understood as a smoothing function, standing for the individual sub-aperture transfer function. The sampling is achieved via the convolution by a Dirac comb of period  $N/h$ . The latter convolution leads to replicating the spectrum in the frequency domain due to the periodic spatial sampling. There is no loss of information due to the sinc function, except at the null points located at multiple integers of  $N/h$  in the Fourier space.

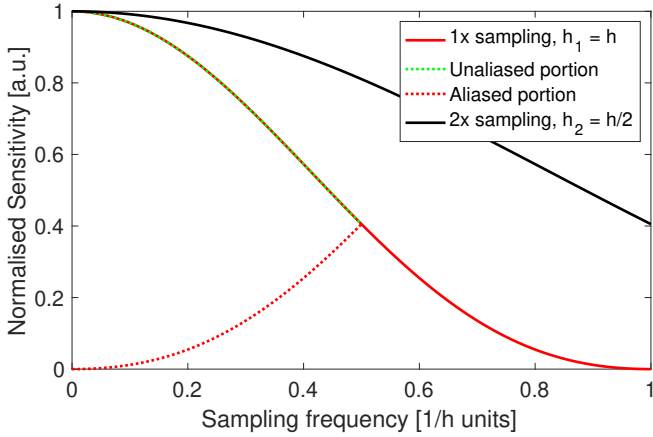
Let us also define the super-resolution factor  $\mathcal{F}$  as the ratio between the maximum spatial frequency,  $k_{\max}$ , that can be reconstructed without aliasing and the native Nyquist frequency  $\nu = 1/2h$ , that is:

$$\mathcal{F} = \frac{k_{\max}}{\nu}. \quad (8)$$

So,  $\mathcal{F}$  stands for the gain in resolution achieved with  $N$  sampling grids, with respect to the case of a single grid with sampling step,  $h$ .

The meta-uniform sampling case (Eq. (7)) carries the following key features: Firstly, from the Nyquist–Shannon Theorem, we know that in the case of Eq. (7), the signal can be reconstructed up to a frequency of  $k_{\max} = N/2h$  without aliasing. The super-resolution factor  $\mathcal{F}$  is therefore equal to  $N$ . Thus, combining properly shifted sampling grids leads to increasing the spatial resolution by a factor,  $\mathcal{F}$ , thus decreasing the aliasing. Secondly, the sub-aperture size is preserved, hence, the sensitivity function remains the same as the native one. Finally, the SR technique offers the capability to design a ‘meta-sensor’ whose sampling





**Fig. 2.** Normalised sensitivity of a virtual averaging-and-sampler sensor used for illustration purposes only. The black curve corresponds to the sensitivity function of a sampling grid of step  $h/2$ . The dotted curve represents the sensitivity of a sampling grid of step  $h$ . The red curve is obtained by dealiasing the spectral sensitivity up to  $1/h$ , that is, the first null of the sinc function with a super-resolution factor of 2; by combining two sampling grids of step  $h$  relatively shifted by  $h/2$  with respect to each other.

step is independent from the sub-aperture size. This independence allows us to decouple the desired spatial sampling period from the sensitivity function in multi-WFS AO systems, which is a new design paradigm.

Considering a super-resolution factor  $\mathcal{F}$  of 2, it is interesting to note that there is no longer any overlap in the spectra – until the first null of the sinc function. Figure 2 highlights the benefit of such a feature in terms of aliasing mitigation.

The sensitivity is limited by the sinc function in the Fourier space (Eq. (7)) and by the aliasing of its replicated self. From Fig. 2, we can see that beating the  $\nu = 1/2h$  spatial resolution limit leads to aliasing reduction when introducing a SR factor of  $\mathcal{F} = 2$  (red curve). However, the sensitivity is greatly reduced with respect to a case with a twice larger native sampling frequency (black curve). With a SR factor of  $\mathcal{F} > 2$ , it would be possible to go beyond a resolution of  $2\nu = 1/h$  (not shown on this figure). Nevertheless, the sensitivity would again be limited by the sinc function that depends only on the sub-aperture size. For spatial frequencies higher than the first null of the sinc function, the sensitivity (red curve) is much lower than what would be achieved with a native sampling that is four times greater. This illustration indicates that noise propagation should be analysed with care in presence of SR. Indeed, on the one hand, the sensitivity is limited by the size of the sub-aperture. On the other hand, a larger sub-aperture size allows a higher signal-to-noise ratio to be reached. These two opposite effects ought to be taken into account to evaluate the noise propagation of a super-resolved system.

### 2.3. Signal reconstruction: Number of sampling grids and optimal offsets

In this section, we wish to address signal reconstruction by analyzing the number of sampling grids required for a reconstruction without aliasing, as a function of their relative spatial offsets. From the Nyquist–Shannon sampling theorem, we know that the choice of  $N$  offsets according to Eq. (4) gives a meta-uniform sampling that provides a SR factor of  $\mathcal{F} = N$ . We analyse here whether or not there are other ways to select the offsets in order

to reach a higher SR factor, or at least to preserve it, without adding extra sampling grids.

In Appendix A, we analyze the reconstruction of band-limited signals by convolutions with sinc interpolation functions within the PNS framework. Under conditions that simplify the analysis (but that may be too restrictive), we show that with  $N$  sampling grids,  $\mathcal{F}$  cannot be higher than  $N$ , independently of the offsets choice. Moreover, the meta-uniform sampling is the only solution with  $N$  sampling grids that reaches  $\mathcal{F} = N$ , meaning that it allows us to reconstruct the signal without aliasing until the frequency  $N\nu$ . With non-uniform configurations, that is, with samples that are not equally spaced, more than  $N$  sampling grids are necessary. Otherwise, any departure from the meta-uniform sampling scheme will generate some reconstruction error. We can show that this error term is bounded and converges to zero when tending towards the meta-uniform sampling. However, we see hereafter that other hypotheses and more elaborated reconstruction approaches allow us to show that some well-chosen non-uniform sampling schemes can reach the same performance as the meta-uniform sampling.

In practice, it may be tricky to place the sampling grids exactly at the positions required to generate a meta-uniform sampling. So, we considered how many sampling grids would be needed to achieve a given SR factor  $\mathcal{F}$ , even if the sampling grids are disposed without particular care, possibly with random offsets. We show that with  $N = 2n_0 - 1$  randomly positioned sampling grids, where  $n_0 = \lceil \mathcal{F} \rceil$  is the ceiling function of  $\mathcal{F}$ , we can always reach a SR factor at least equal to  $\mathcal{F}$ .

To be specific and give some concrete examples, if we want to reconstruct a signal whose maximal frequency satisfies  $\nu < k_{\max} \leq 2\nu$ , we need  $1 < \mathcal{F} \leq 2$ , and  $n_0 = 2$ . So, in the mono-dimensional case, we may use two sampling grids providing a meta-uniform sampling, or three sampling grids that are randomly positioned, excluding the trivial case where some offsets are multiple of the native sampling step. In the bi-dimensional case, we can show that four sampling grids are needed with a meta-uniform sampling scheme, while nine randomly positioned sampling grids are sufficient.

As previously mentioned, with a more elaborated approach based on Lagrange polynomials, it is possible to reconstruct non-uniformly sampled signals without aliasing, even outside the PNS framework. Indeed, the Paley–Wiener–Levinson theorem (Marvasti 2001; Wiener 1930; Paley & Wiener 1934; Levinson 1940), which generalizes the Whittaker–Shannon–Kotelnikov sampling theorem (Marvasti 2001) from uniform to non-uniform samples, states that a band-limited signal can be reconstructed from its samples if the two following sufficient conditions are respected.

Firstly, according to Beutler (1966), the average sampling rate  $k_e$ , that is the inverse of the average sampling step  $h_e$ , satisfies the generalized Nyquist–Shannon condition:

$$k_e = \frac{1}{h_e} \geq 2k_{\max}. \quad (9)$$

In the PNS framework,  $h_e$  is always equal to  $h/N$ . To reach a SR factor  $\mathcal{F} = N$ , at least  $N$  sampling grids are necessary.

Secondly, the Kadec condition (Kadec 1964) states that the non-uniform sampling should be constructed by relocating the points of a uniform sampling of step  $1/2k_{\max}$  no further than  $1/8k_{\max}$  from their original location.

We have seen that under some conditions, both uniform and non-uniform sampling enable a theoretically perfect reconstruction of the sampled band-limited signal. In the PNS framework,

the  $N$  sampling grids' offsets shall be close to but not necessarily equal to a multiple of  $h/N$ , in order to achieve optimal SR ( $\mathcal{F} = N$ ). We show in Sect. 3.3 that numerical simulations confirm the results from the first part of this paper. The meta-uniform sampling is indeed optimal, in the sense that it provides the best reconstruction performance. Nevertheless, it is important to note that the performance is improved as soon as any non-zero shift is introduced. Moreover, there is a shift range around the optimal shift that provides stable performance, which indicates its robustness. This shift range is consistent with the Kadec condition (Kadec 1964) for a perfect reconstruction with non-uniform sampling.

### 3. Super-resolution wavefront reconstruction with Shack-Hartmann WFSs

#### 3.1. Bi-dimensional model

We go on to transpose the previous models to a SH-WFS, whose measurements  $\mathbf{s}(\mathbf{x})$  are well approximated by the geometrical-optics linear model (Rigaut et al. 1998):

$$\mathbf{s}(\mathbf{x}) = \mathcal{G}\boldsymbol{\varphi}(\mathbf{x}) + \boldsymbol{\eta}(\mathbf{x}), \quad (10)$$

where  $\mathcal{G}$  is a phase-to-slopes linear operator mapping aperture-plane guide-star wavefronts  $\boldsymbol{\varphi}(\mathbf{x})$  onto WFS measurements over a bi-dimensional space indexed by  $\mathbf{x} = [x, y]$ ;  $\boldsymbol{\eta}(\mathbf{x})$  represents white noise due to photon statistics, detector read noise and background photons. Both  $\boldsymbol{\varphi}$  and  $\boldsymbol{\eta}$  are zero-mean functions of Gaussian probability distributions and known covariance matrices  $\Sigma_{\boldsymbol{\varphi}}$  and  $\Sigma_{\boldsymbol{\eta}}$  respectively. Noise is assumed both temporally and spatially uncorrelated.

It can be shown for the nominal case when the sub-aperture size  $d$  coincides with the sampling step  $h$  (Correia et al. 2017) that:

$$\mathcal{G} = \text{III}\left(\frac{\mathbf{x}}{h}\right) \times \left[ \Pi\left(\frac{\mathbf{x} - 1/2}{d}\right) \otimes \nabla \right], \quad (11)$$

where  $\otimes$  is a two-dimensional (2D) convolution product,  $\times$  is a point-wise multiplication,  $\Pi(\cdot)$  is the top-hat separable function,  $\nabla$  is the gradient operator, and  $\text{III}(\mathbf{x})$  is a comb function (a bi-dimensional sum of Dirac delta functions) that represents the sampling process. In the following, we consider the classical case of  $d = h$ .

Now the following argument stands: if we properly combine four sets of measurements provided by convolutional operators:

$$\mathcal{G} = \begin{bmatrix} \mathcal{G}_1 \\ \mathcal{G}_2 \\ \mathcal{G}_3 \\ \mathcal{G}_4 \end{bmatrix} \quad (12)$$

with a relative offset between them of  $h/2$ , corresponding to the meta-uniform sampling in 2D, that is:

$$\begin{aligned} \mathcal{G}_1 &= \text{III}\left(\frac{\mathbf{x}}{h}\right) \times \left[ \Pi\left(\frac{\mathbf{x} - [1/2, 1/2]}{h}\right) \otimes \nabla \right], \\ \mathcal{G}_2 &= \text{III}\left(\frac{\mathbf{x} - [0, 1/2]}{h}\right) \times \left[ \Pi\left(\frac{\mathbf{x} - [1/2, 1/2]}{h}\right) \otimes \nabla \right], \\ \mathcal{G}_3 &= \text{III}\left(\frac{\mathbf{x} - [1/2, 0]}{h}\right) \times \left[ \Pi\left(\frac{\mathbf{x} - [1/2, 1/2]}{h}\right) \otimes \nabla \right], \\ \mathcal{G}_4 &= \text{III}\left(\frac{\mathbf{x} - [1/2, 1/2]}{h}\right) \times \left[ \Pi\left(\frac{\mathbf{x} - [1/2, 1/2]}{h}\right) \otimes \nabla \right], \end{aligned} \quad (13)$$

it leads to a meta-SH system with half the sampling step and the same averaging cell size or sub-aperture width as in the native case:

$$\tilde{\mathcal{G}} = \text{III}\left(\frac{\mathbf{x}}{h/2}\right) \times \left[ \Pi\left(\frac{\mathbf{x} - 1/2}{h}\right) \otimes \nabla \right], \quad (14)$$

therefore lifting the sub-aperture imposed spatial cut-off frequency, which becomes, in the case of a SR factor of  $\mathcal{F} = 2$ :

$$k_{\max} = 2\nu = \frac{1}{h}. \quad (15)$$

Based on the fact that the SH-WFS measurement model is a set of convolution integrals, the treatment in the spatial-frequency domain is straightforward. Let the Fourier-domain representation of Eq. (10):

$$\tilde{\mathbf{s}}(\boldsymbol{\kappa}) = \tilde{\mathcal{G}}\tilde{\boldsymbol{\varphi}}(\boldsymbol{\kappa}) + \tilde{\boldsymbol{\eta}}(\boldsymbol{\kappa}), \quad (16)$$

with  $\boldsymbol{\kappa} = (k_x, k_y) \in \mathbb{R}^2$  as the frequency vector and symbol  $\tilde{\cdot}$  used for Fourier-transformed variables. The time-dependence could be added in with recourse to the frozen-flow hypothesis. Using common transform pairs for the individual operations (Oppenheim & Willsky 1997), the Fourier representation of the measurements in Eq. (10) given in Eq. (16) can be expanded to:

$$\begin{aligned} \tilde{\mathbf{s}}(\boldsymbol{\kappa}) &= 2i\pi h \sum_{\mathbf{m}} (\boldsymbol{\kappa} + \mathbf{m}/(h/N)) \tilde{\boldsymbol{\varphi}}(\boldsymbol{\kappa} + \mathbf{m}/(h/N)) \\ &\quad \times \tilde{\Pi}(\boldsymbol{\kappa}h + \mathbf{m}) \times e^{i\pi(\boldsymbol{\kappa}h + \mathbf{m})} + \tilde{\boldsymbol{\eta}}(\boldsymbol{\kappa}). \end{aligned} \quad (17)$$

From Eq. (17), it is apparent that the wavefront sensing operator  $\tilde{\mathcal{G}}$  is not purely a spatial filter except for functions within the SH-WFS pass-band (Ellerbroek 2005).

#### 3.2. Trade-off between sensitivity and noise propagation

In practical terms, any interested AO practitioner may wonder (quite rightly) whether SR signal reconstruction above the nominal Nyquist–Shannon cut-off is associated with a somewhat greater noise propagation, which could limit the prospective SR advantages. We first address this key point in the spatial-frequency domain and later using the Karhunen–Loeve (KL) modal basis. We show in either case that SR grants comparable noise propagation without sacrificing effective sensitivity.

##### 3.2.1. Sensitivity functions revisited: Fourier S/N

We start by revisiting commonly-accepted expressions for sensitivity functions, often called the Fourier signal-to-noise ratio (S/N; Verinaud 2004). The sensitivity defined as a S/N (the ratio between the produced signal and the noise affecting the measurement) is a function of the effective number of photons and spot size on each sub-aperture, as follows

$$S/N = \Xi_{\text{SH}}^2 / \sigma_{\text{ph}}^2. \quad (18)$$

The SH-WFS sensitivity can be readily computed from Eq. (17) for functions in the pass-band assuming the compact form from Correia & Teixeira (2014):

$$\Xi_{\text{SH}}^2 = (2i\pi h k_x)^2 \left( \frac{\sin(\pi k_x)}{\pi k_x h} \right)^2 = \Xi_{\text{grad}}^2 \Xi_{\text{aver}}^2, \quad (19)$$

where we assume the effective SH signal to be the angle of arrival on each sub-aperture instead of the phase difference at opposite sub-aperture edges (there is a normalisation factor,  $h$ , between both). The 2D case is now straightforward to obtain. Let the following photon and detector noise expressions (Roddiier 1999; Hardy 1998):

$$\sigma_{\alpha, \text{ph}}^2 = \frac{1}{8} \frac{1}{n_{\text{ph}}} \theta_{\text{spot}}^2, \quad (20)$$

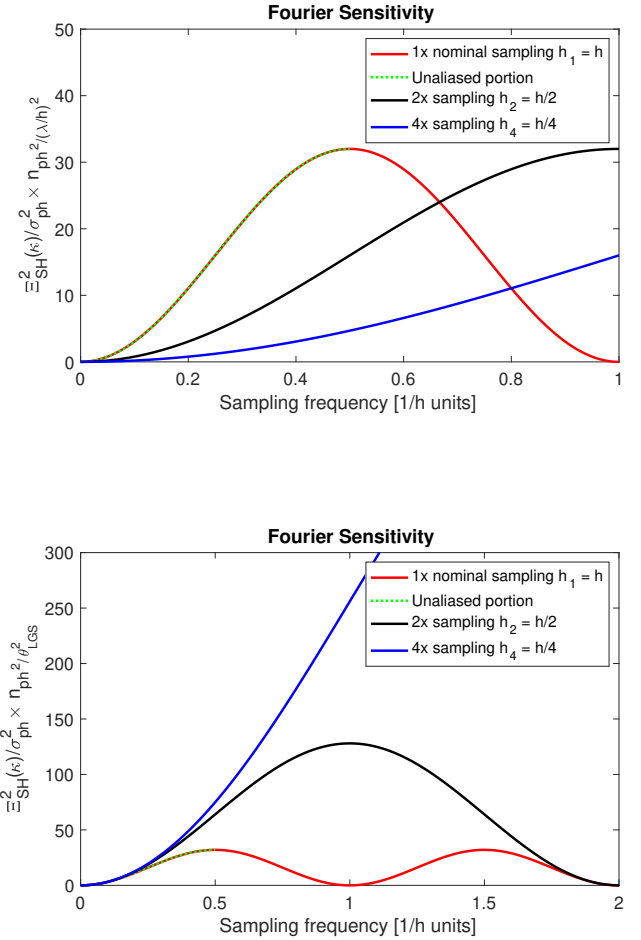
$$\sigma_{\alpha, \text{det}}^2 = \frac{1}{12} \frac{\sigma_e^2}{n_{\text{ph}}^2} \left( \frac{N_s^2}{N_D} \right)^2 \theta_{\text{spot}}^2, \quad (21)$$

with  $n_{\text{ph}}$  as the number of incident photons per frame and per sub-aperture,  $\theta_{\text{spot}}$  the spot size,  $\sigma_e$  the rms number of detector read-out photo-electrons, and  $N_s$  and  $N_t$  as the window used for processing and sub-aperture diffraction sizes respectively. In the remainder of this paper, we limit the analysis to photon noise only, since it is not far-fetched to consider practically zero-noise detectors today.

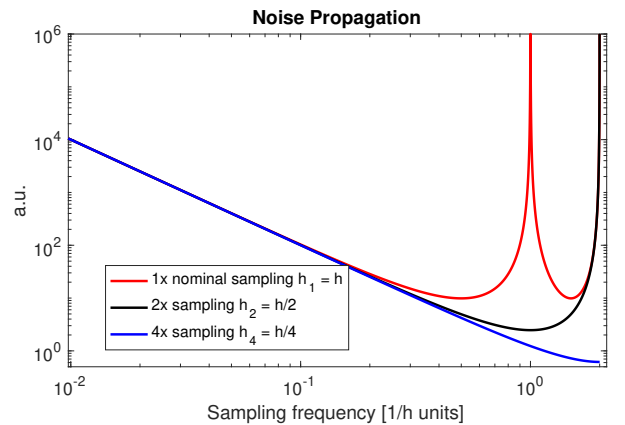
In diffraction-limited cases with point-sources, then  $\theta_{\text{spot}} = \frac{\lambda}{h}$ . It is straightforward to check that if we chose to split the aperture into  $n_{\text{SA}}$ -times as many sub-apertures (linearly across the pupil), the photon-noise is increased by a factor of  $n_{\text{SA}}^4$  that results from the product of  $n_{\text{SA}}^2$  fewer photons per sub-aperture and a diffraction spot size that is twice as large, resulting in another  $n_{\text{SA}}^2$  factor for the bi-dimensional case. Considering all the measurements over the whole pupil, an average over  $n_{\text{SA}}^2$  times more measurement points is performed, resulting in a final  $n_{\text{SA}}^2$  sensitivity decrease factor. This factor is at the origin of the so-called ‘full-aperture gain’, which leads to the use of a single full aperture spot measurement for tip and tilt modes. A comparative plot shown in the top panel of Fig. 3 illustrates this fact. This would lead to the conclusion that using a super-resolved meta-SH-WFS setting leads to sensitivity gains over a twice-resolved SH and even to a four-times-resolved SH over the original sampling band  $[0 : \frac{1}{2h}]$  and some portions of the  $[\frac{1}{2h} : \frac{1}{h}]$ . Therefore, SR appears very well suited for AO applications with diffraction limited WFS, for instance, those sensing with a corrected near-infrared natural guide star (NGS).

We do, however, expect to deploy multi-WFS mostly on laser-tomographic AO cases. In this case, splitting the aperture into fewer sub-apertures leads to small (or none at all) sensitivity gains because the spot is no longer a function of the diffraction but, instead, it is of a fixed size. In the bottom panel of Fig. 3, we can inspect the relative sensitivities: they are roughly the same until half the control radius of the nominal case of  $[0 : \frac{1}{4h}]$ ; by this time, the larger sub-aperture size averaging function (the sinus cardinal) starts bending down the sensitivity curve. In the SR case, we can only be as sensitive as the red curve, yet with the aliasing cut-off frequency being a function of the over-sampling used. The periodic nature of the sensitivity function with an initial zero at  $\kappa = 1/h \text{ m}^{-1}$  tells us that a SR factor of up to two is of practical relevance but not much beyond. Indeed, the high noise amplification at the frequencies with close-to-zero sensitivity is to be traded-off for increased sensitivity at frequencies located beyond the native cut-off frequency.

If we opt to evaluate the noise propagation instead, we can see from Fig. 4 that SR would not lead to higher figures at low spatial-frequencies and that it would extend the effective maximum reconstructed frequency avoiding the zeros of the sinc<sup>2</sup> function for which the noise propagation is infinite. The sensitivity evaluated on a single sub-aperture tells us little about how



**Fig. 3.** Fourier S/N functions. *Top:* Fourier S/N for a diffraction-limited wavefront sensing case (Eq. (6)). *Bottom:* seeing-limited or laser spot-sensing case. The number of incident photons is considered the same for all configurations. The *bottom* panel assumes a spot size with one unit.



**Fig. 4.** Noise propagation coefficient as a function of spatial frequency. This plots are computed as the inverse of the S/N functions from Fig. 3. The fluxes are scaled by the collecting area, namely, the square of the sub-aperture width.

it is distributed over modes spanning the aperture. For tip-tilt, it is easy to compute since the global estimate is a straight average of each sub-aperture’s measurements, resulting in an effective  $1/n_{\text{SA}}^2$  gain (diffraction-limited sensing) or no gain (seeing or LGS spot limited sensing). In the next section, we present a more general case of modal reconstruction for multi-WFS AO systems.

**Table 1.** Simulation configurations.

Telescope diameter	8 m
Number of WFSs	4 SH-WFS
Native resolution case	$n_{SA} = 40$ sub-apertures
Lower resolution cases	$n_{SA} = 20$ or $10$ sub-apertures
Turbulent wavefront	1 layer in the pupil plane: KL modes

**Notes.** Four SH-WFS grids are either co-aligned as in a classical tomographic AO system or shifted with respect to each other. The super-resolution sampling scheme is shown in Fig. 5 case (b) for a SR factor of 2, i.e. four WFS in 2D.

### 3.2.2. Noise propagation: modal analysis

In order to properly treat the general case, the noise propagation is considered on a per-mode basis, following Rigaut & Gendron (1992):

$$\text{n.p.} = \text{diag}\{RR^T\}, \quad (22)$$

where  $R$  stands for the reconstructor matrix. We use this formulation instead of the original one provided by Eq. (14) in Rigaut & Gendron (1992) due to the use of regularised reconstruction. If the least-squares solution were adopted, we would find Rigaut's  $(D^T D)^{-1}$ , with  $D$  the system modal interaction matrix concatenating the sensor's response to each mode.

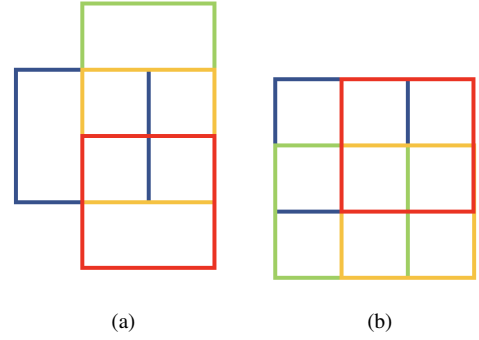
Figure 4, which is essentially the inverse of Fig. 3 as can readily be seen, depicts the noise propagated as a function of spatial frequency. When accounting for the sub-aperture size in the computation of the noise propagation, we can confirm that SR configurations do not propagate more noise than classical configurations, along lower spatial frequencies.

### 3.3. Numerical application to multi-WFS-AO systems

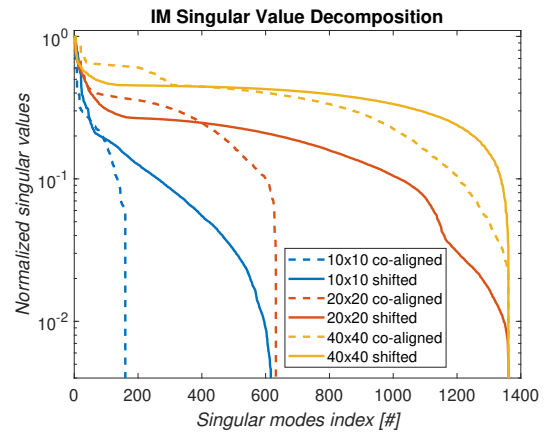
In this section, we numerically simulate the SR concept applied to a multi-SH-WFS AO system. First of all, we can easily confirm that SR can be promoted by building a multi-SH-WFS model and measuring the Interaction Matrix (IM) between the phase and the meta-SH-WFS, combining several LR WFS, within the configuration described in Table 1.

We go on to compute a set of modal IMs, between the 1364 KL modes as input and the four SH-WFS slopes as output, this for all configurations ( $40 \times 40$ ,  $20 \times 20$ ,  $10 \times 10$ , co-aligned, or shifted). The rank of the IM indicates the number of degrees of freedom that the super-resolved system is sensitive to. The singular value decomposition (SVD) of the IM highlights the sensitivity of the system to the singular modes. In other words, the singular value associated to each singular mode indicates how well such a mode is seen by the system. Figure 6 highlights the benefit provided by the shifted configurations. In the classical cases (solid lines), the sensitivity drops to 0 for a KL index equal to the number of slopes measured by each individual SH-WFS. On the other hand, with SR (dashed lines), the sensitivity is enhanced and the sensitivity limit is pushed towards that of the co-aligned cases with double native resolution. Concretely, the super-resolved  $10 \times 10$  (resp.  $20 \times 20$ ) configuration provides some additional sensitivity up to the singular modes located at the sensitivity limit for the  $20 \times 20$  (resp.  $40 \times 40$ ) co-aligned configuration.

We then perform an open-loop reconstruction of the phase by inverting the previously computed IMs in a MMSE fashion. We note that each KL mode past the LR SH-WFS cut-off frequency



**Fig. 5.** SR configurations illustrated via a projection in the pupil plane of one sub-aperture from each WFS (1 color for each): (a) 2D surface is improperly paved with shifts along the sampling grid's axes; (b) 2D surface is properly paved with the addition of shifts at  $45^\circ$  with respect to the sampling grid's axes, as in Eq. (13).

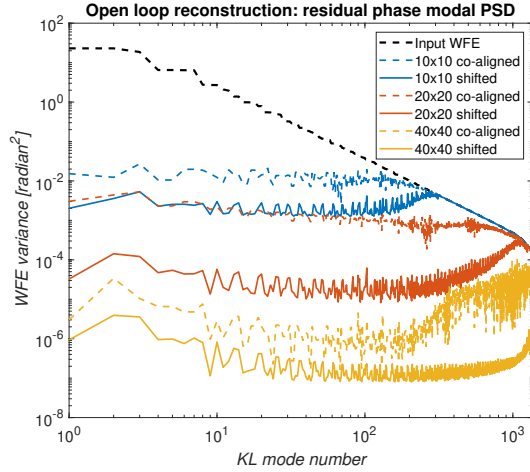


**Fig. 6.** Singular value decomposition of a multi-WFS mono layer modal IM. The singular values are plotted as a function of the singular modes index. The solid lines corresponds to 4 co-aligned SH-WFS while the dashed lines represent the shifted configurations shown in Fig. 5b.

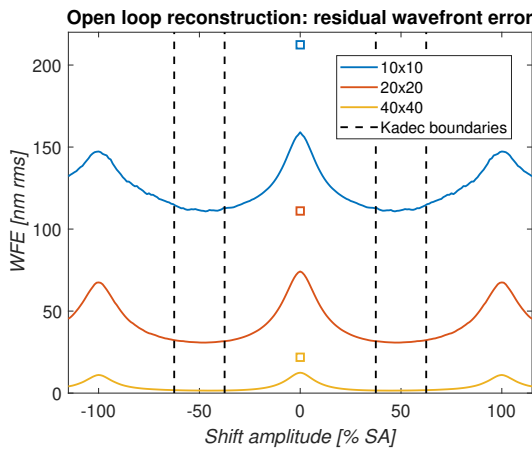
$\nu = 1/(2h)$  aliases onto its corresponding low-frequency spectrum, yet the inversion of the meta SH-WFS IM plays the role of a numerical de-scrambler as if the samples were spatially reordered and, therefore, de-aliased. Figure 7 indicates that the SR configurations significantly reduce the reconstruction error with respect to the co-aligned cases. The cut-off frequency is twice higher than the respective classical case with as many reconstructed modes as number of sensing element per WFS  $\frac{\pi}{4} n_{SA}^2$ . Moreover, the aliasing features vanish close to the cut-off frequency for the super-resolved cases. Recent end-to-end closed-loop simulations performed in the context of ELT tomographic AO systems (MORFEO and HARMONI LTAO) confirm the benefits of the SR technique in terms of aliasing mitigation at an iso number of sub-apertures, as well as the potential reduction of the number of sub-apertures, while preserving a similar level of performance as the classically dimensionned AO system (Fusco et al. 2022). Summing the modal residual variances shown in Fig. 7 provides an estimation of the total reconstruction wavefront error (WFE) for the shifted cases.

In Fig. 8, we evaluate the residual WFE as a function of shift amplitude in percentage of a subaperture width. A first observation is the performance gain that can be obtained when reconstructing four times more modes than the number of degrees of freedom of a single WFS (colored squares). It can be explained by the 'statistical SR' properties. In this case, the MMSE process extrapolates the measurements statistically towards larger spatial





**Fig. 7.** Open-loop reconstruction of an incident phase corresponding to a Kolmogorov turbulence with  $r_0 = 12.6$  cm. The dashed lines represent the co-aligned cases. The solid lines stand for the super-resolved cases achieved by combining four SH WFS that are properly shifted by half ( $1/2$ ) of a sub-aperture. The same number of modes is reconstructed for the co-aligned cases and their respective shifted cases, that is, up to  $4 \cdot \frac{\pi}{4} n_{SA}^2$  modes, which is the spatial resolution that would be provided by a single WFS with twice as many sub-apertures per diameter.



**Fig. 8.** Residual wavefront error from open-loop reconstruction as a function of shift amplitude for the shifted and super-resolved configurations. In each case, up to four times the native number of modes ( $4 \cdot \frac{\pi}{4} n_{SA}^2$ ) are reconstructed. The dashed black lines represent the Kadec condition boundaries within which the non-uniform samples can be perfectly reconstructed. The colored squares mark the reconstruction error from the co-aligned case when  $\frac{\pi}{4} n_{SA}^2$  modes are reconstructed.

frequencies, by assuming the power spectrum described in the Kolmogorov covariance priors. Next, we focus on the ‘geometric SR’ properties. We can observe a periodicity of the achieved WFE with a periodic shift that is a multiple integer of the sub-aperture width, as expected. However, with any other fractional shift amplitude, the WFE is improved with again an optimal value of one-half ( $1/2$ ) of a sub-aperture, which corresponds to a meta-uniform sampling generated by the super-resolved system. At this optimal offset of half of a sub-aperture width, it is interesting to note that the  $10 \times 10$  (resp.  $20 \times 20$ ) shifted configuration achieves a WFE that tends towards that from a  $20 \times 20$  (resp.  $40 \times 40$ ) co-aligned system with a number of reconstructed modes equal to the native number of degrees of freedom ( $\frac{\pi}{4} n_{SA}^2$ ).

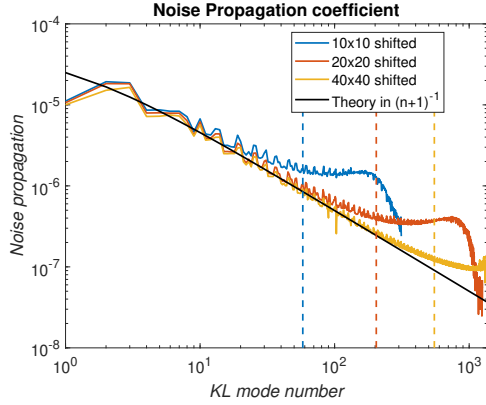
As already pointed out, the ‘geometric SR’ provides an improvement of the WFE as soon as there is a small offset introduced. So, even non-optimal shifts induce a dramatic improvement of the reconstruction quality. Furthermore, the stability of the WFE around the optimal shift values is remarkable. In summary, Fig. 8 confirms on one hand that the meta-uniform sampling is optimal, as then demonstrated in Appendix A. On the other hand, it appears unnecessary to accurately tune the system with such optimal offsets since the performance appears weakly sensitive to the offset value. Actually, this observation is consistent with the Kadec condition (Kadec 1964), which corresponds to a range of shifts of  $\pm 12.5\%$  of a sub-aperture width around the  $1/2$  sub-aperture offset, within which the samples can be perfectly reconstructed. In our simulation, the MMSE reconstructor was not optimized for each shift value, hence, there is a slight performance degradation when going away from the meta-uniform sampling case. The Kadec range can be interpreted as a tolerance to uncontrolled misalignments, as long as these ‘mis-registrations’ are calibrated or at least known. Such feature opens the door to a diversity of geometrical transformation breaking the symmetries of the wavefront sampling topology. Furthermore, Agapito et al. (2020) has suggested that SR is more robust to model calibration errors than ‘classical’ AO system of the same dimension.

In the case of a multi-WFS tomographic AO system, super resolution is ‘built-in’ when it comes to the turbulent layers conjugated in altitude. Indeed, when projecting the lenslet array grids onto the altitude layers, the off-axis WFS grids are shifted proportionally to the conjugation altitude. It is straightforward to assume that there are some altitude layers where the resulting shift is a multiple integer of the sampling step, hence, no super resolution is granted in this case. An optimized geometrical arrangement, possibly including a rotation of the lenslet array grids with respect to each other, would allow for this effect to be mitigated.

Finally, we evaluate how SR behaves in terms of noise propagation. Figure 9 (essentially a numerical confirmation of Fig. 4) depicts the noise propagated as a function of the KL mode number for a combination of four WFS that are: (1) co-aligned and (2) optimally shifted by half a sub-aperture width of SH-WFS. The noise propagation follows the expected  $(n + 1)^{-1}$  power law with the shifted system (on which SR is applied) not exceeding the co-aligned case. We note that for each of the tested cases, the noise propagation within the original cut-off frequency range is almost indistinguishable from that of the co-aligned system; whereas when it goes past the original Nyquist–Shannon cut-off frequency, it rolls off, asymptotically saturating to the  $(n + 1)^0$  regime.

#### 4. SR WFR with pyramid WFSs

We argue that the pyramid wavefront sensor (P-WFS) also benefits from an effect that is comparable to SR in the Shack-Hartmann WFS (SH-WFS). Unlike the latter, the SR principle can be directly applied to a single P-WFS when we allow a more general formalism than Eq. (5), whereby the samples are obtained from transformations of the initial function. In the case of the P-WFS, each quadrant applies a different transformation, each of which is required for proper signal reconstruction. However, two diagonally opposite quadrants measure the same information for the spatial frequencies larger than the modulation radius (Guyon 2005). Thus, the mono-dimensional SR formalism presented in this paper (see Sect. 2) can, in principle, be applied to the corresponding modes for each pair of quadrants



**Fig. 9.** Noise propagation coefficients, computed as  $\text{diag}\{RR^T\}$ , shown in logarithmic scale and scaled by the flux collecting surface of the sub-aperture. The shifted cases are plotted with solid lines. For the sake of visibility, dashed lines are overlayed at the level of the spatial frequencies where the aliasing starts being prominent for the co-aligned cases. Along the low order modes, the noise propagation is quasi identical for all cases and fit the expected law in  $(n+1)^{-1}$ . Then, when exceeding the original Nyquist–Shannon cut-off frequency, the noise propagation coefficient converges towards a constant asymptotic limit.

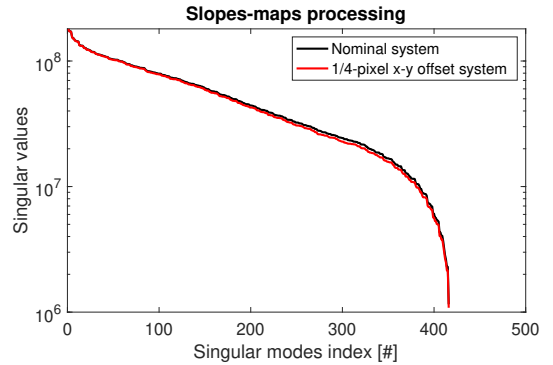
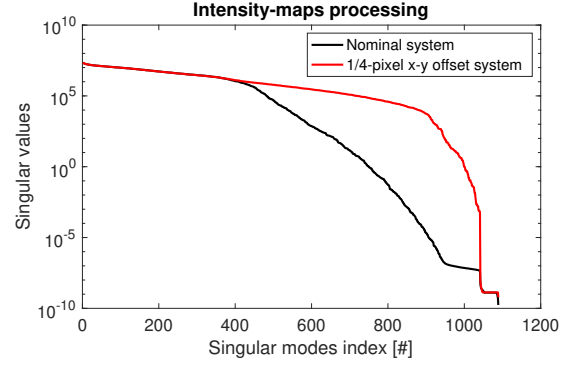
whose pixel grids are shifted in one direction by a fraction of a pixel.

Here, the multiple samples are nominally four diffraction patterns produced at the corresponding re-imaged pupil planes, however, with the sampling grid pixels shifted with respect to one another in such a way that their sampling becomes interlaced by a fractional pixel. By adjusting the P-WFS apex angles, one such configuration can be readily obtained. The same process can, in principle, be carried out simply by rotating the pyramid optic with respect to the detector. This relaxes the specification and manufacturing constraints, yet in order to obtain the optimal SR-enabling configuration, we may end up re-introducing anew such design constraints that we wanted to relax in the first place. Clearly, more investigation on the optimality and robustness are required and we invite the community to contribute to this endeavour.

To exploit the SR enhancement however, the use of the four diffraction intensity patterns (i.e. intensity maps) instead of the customary slope maps appears to be required. It is known that the ‘x’ and ‘y’ slopes-maps contain the complete information on perfectly aligned systems, that is, exactly registered quadrants and a perfect four-sided prism with homogeneous and stationary pupil illumination (Deo et al. 2018). However, additional slope maps are required when these conditions fail. In our opinion, the ‘extension’ proposed in Deo et al. (2018) is unnecessary since it represents a full-rank linear combination of diffraction intensity patterns. This intermediate step adds nothing to the use of intensity maps directly, as it neither adds nor removes useful information to that originally contained in the pixels’ intensities directly. Furthermore, it can only be detrimental to the pipelined pixel processing by creating additional idle time before matching pixels become available during detector read-out.

Mathematically, out of the eight possible linear combinations (we can take each intensity map either positive or negative and there are four of these), one full-rank slopes-map is:

$$\begin{bmatrix} s_x \\ s_y \\ s_{xy} \\ s_t \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \\ i_4 \end{bmatrix} / F_t, \quad (23)$$



**Fig. 10.** SVD decomposition on a  $16 \times 16$  P-WFS system with a  $33 \times 33$  DM. *Top*: intensity map processing. *Bottom*: slopes-maps processing. We clearly see that using the intensity-maps on an offset pixel sampling system leads to sensitivity gains, whereas the same is not true if the slopes-maps were to be processed.

where  $F_t$  is a scalar value that represents the total flux captured on all the four re-imaged pupils. We have made a slight change to the last entry,  $s_t$ , which in Deo et al. (2018) is the total flux  $s_t = [1, 1, 1, 1] [i_1, i_2, i_3, i_4]^T$  but leads essentially to the same results for the rank of either transformation matrix is 4. As a matter of fact, we do go beyond the recommendations of Deo et al. (2018) in that we do not restrict ourselves to developing a misalignment mitigation strategy. Under the SR framework, we claim that by expressly misaligning the PWFS system, we can enhance sensitivity and therefore gain in the delivered optical performance.

The use of intensity-maps has obvious implications for real-time processing (i.e. twice as large reconstruction matrix) that needs to be carefully counter-balanced against the performance gains. To illustrate this behaviour, Fig. 10 shows the SVD decomposition from a  $16 \times 16$  lenslet based P-WFS system with a  $33 \times 33$  regular DM. We clearly see that the introduction of a relative offset (translation) in the pixels sampling grid leads to sensitivity gains provided that the intensity-maps are processed. The opposite is obtained when the slopes-maps are processed instead. The mathematical demonstration and interpretation will be provided in a forthcoming paper.

## 5. Conclusion and outlook

Super-resolution is a computational imaging technique that is aimed at upscaling a signal’s spatial resolution from multiple lower resolution samples. Wavefront reconstruction – a process at the very heart of astronomical adaptive optics systems –

provides a particularly fertile setting for the application of SR. Unlike in imaging systems where the ultimate spatial resolution at the focal plane hits the fundamental diffraction limit, operating in the wavefront domain in or near the pupil plane offers the potential to significantly increase the sensing resolution, within the technological ability to sample the wavefront (spatially and temporally) subject to the inescapable concerns around adequate S/N.

To make SR viable, we have identified several ways to introduce unique phase signatures in the transverse plane to the beam propagation, the simpler one being a straight offset. As it stands, SR promises to lift instrument design assumptions, for instance, that  $n \times n$  measurement samples are needed to drive a  $n \times n$  actuators DM, that the registration between the DM and the WFS sampling grid needs to be finely controlled, and that the guide-star asterism be regular (along with a few other aspects), leading to a major overhaul in how the AO community may approach the designs of future instruments. By decoupling the WFS sampling step from the averaging cell size (sub-aperture width), we are led to a paradigm shift where the same AO control performance can be achieved with fewer WFS sampling points. This, in turn, leads to fewer detector pixels, smaller systems and ultimately a lighter computational load. Another strategy is to optimize the performance while keeping the same number of sensing degrees of freedom, by reconstructing modes beyond the cut-off frequency and mitigating the aliasing. Several tomographic AO systems in operation (GALACSI NFM) or in the design phase (MORFEO, HARMONI LTAO, MAVIS, KAPA, GMT's GLAO and LTAO, etc.) could benefit from the implementation of SR in one way or the other.

With  $N$  regular sampling grids of step  $h$ , the optimal SR sampling scheme is obtained by introducing fractional offsets that are multiple of  $\frac{h}{\sqrt{N}}$  in the bi-dimensional case. In this way, we can generate a ‘meta sampler’ corresponding to the so-called ‘meta-uniform’ sampling scheme. However, the non-uniform sampling case is appealing because it relaxes the absolute alignment constraints. Concretely, the location of the grids is flexible. The optimal reconstruction performance can even be reached within the Kadec condition, that is, the departure from the meta-uniform sampling by less than a quarter of the Nyquist-Shannon period. Furthermore, we are yet to establish optimal, or at least practical, SR sampling geometries for future tomographic AO systems with a non-square number of WFSs. With six WFS (MORFEO, HARMONI LTAO) or eight WFS (MAVIS), the solution is not obvious. Preliminary geometry optimization results can be found in Cranney et al. (2021). Some other ways to break the system symmetries are proposed in Appendix B.

We have shown that SR does not entail larger noise propagation coefficients and that is applicable to multi-Shack–Hartmann WFS systems as well as to single pyramid WFS. In either case, the relative alignment accuracy is traded-off for the precision with which the registration is known to inform the model-based SR reconstruction. We do not exclude – in fact, we very much welcome – data-driven, machine-learning algorithms that can further optimize this trade-off. The very fast developments of these techniques in the field of SR applied to imaging may be inspiring. A review of deep learning approaches for imaging SR is, for example, provided in Wang et al. (2021).

**Acknowledgements.** This work benefited from the support of the WOLF & APPLY project ANR-18-CE31-0018 & ANR-19-CE31-0011 of the French National Research Agency (ANR), the project Fortalecimiento del Sistema de Investigación e Innovación de la Universidad de Valparaíso UVA20993 and the project ANID/REDES 190071. It has also been prepared as part of the activities of

ORP H2020 Framework Programme of the European Commission's (Grant number 101004719). Authors are acknowledging the support by the Action Spécifique Haute Résolution Angulaire (ASHRA) of CNRS/INSU co-funded by CNES, the support of LABEX FOCUS ANR-11-LABX-0013 and the support of the programme d'investissement avenir ANR-21-ESRE-0008. We would like to thank Christophe Verinaud and Guido Agapito for reading parts of the manuscript and providing pertinent comments. Finally, we wish to acknowledge the fruitful discussions with Pierre-Yves Madec, Michel Tallon, Miska Le Louarn and Jason Spyromilio at the inception of this idea back in 2016. Thank you for your critical review and wise feedback.

## References

- Agapito, G., Plantet, C., Busoni, L., et al. 2020, *SPIE*, **11448**, 114482S
- Bara, S., Arines, J., & Pailos, E. 2013, *Opt. Eng.*, **53**, 061703
- Beutler, F. J. 1966, *Siam Rev.*, **8**, 328
- Busoni, L., Agapito, G., Plantet, C., et al. 2019, in *6th International Conference on Adaptive Optics for Extremely Large Telescopes, AO4ELT 2019*, Québec, Canada
- Conan, R. 2014, in *Fast Iterative Optimal Estimation of Turbulence Wavefronts with Recursive Block Toeplitz Covariance Matrix*, 9148, 91480R
- Correia, C. M., & Teixeira, J. 2014, *J. Opt. Soc. Am. A*, **31**, 2763
- Correia, C. M., Bond, C. Z., Sauvage, J.-F., et al. 2017, *J. Opt. Soc. Am. A*, **34**, 1877
- Cranney, J., Guihot, A., De Dona, J., & Rigaut, F. 2021, in *2021 Australian New Zealand Control Conference (ANZCC)*, 24
- Deo, V., Gendron, É., Rousset, G., et al. 2018, *A&A*, **619**, A56
- Ellerbroek, B. L. 2005, *J. Opt. Soc. Am. A*, **22**, 310
- Ellerbroek, B. L., & Rigaut, F. 2001, *J. Opt. Soc. Am. A*, **18**, 2539
- Fruchter, A. S., & Hook, R. N. 2002, *PASP*, **114**, 144
- Fusco, T., Conan, J.-M., Rousset, G., Mugnier, L. M., & Michau, V. 2001, *J. Opt. Soc. Am. A*, **18**, 2527
- Fusco, T., Agapito, G., Neichel, B., et al. 2022, *J. Astron. Telescopes Instrum. Syst.*, **8**, 021514
- Gerchberg, R. 1974, *J. Mod. Opt.*, **21**, 709
- Guyon, O. 2005, *ApJ*, **629**, 592
- Hardy, J. W. 1998, *Adaptive Optics for Astronomical Telescopes* (New York: Oxford)
- Irani, M., & Peleg, S. 1990, in *10th IAPR International Conference on Pattern Recognition, Conference C: image, speech, and signal processing, and Conference D: computer architecture for vision in pattern recognition, ICPR 1990*, Atlantic City, NJ, USA, 16–21 June, 1990, 2 (IEEE), 115
- Kadec, M. I. 1964, *Sov. Math. Dokl.*, **5**, 559
- Kohlenberg, A. 1953, *J. Appl. Phys.*, **24**, 1432
- Levinson, N. 1940, in *American Mathematical Society Colloquium*, 26
- Marvasti, F. 2001, *Nonuniform Sampling: Theory and Practice*, 1, 132
- Neichel, B., Fusco, T., & Conan, J.-M. 2009, *J. Opt. Soc. Am. A*, **26**, 219
- Oberti, S., Le Louarn, M., Dialaiti, E., et al. 2017, MAORY design trade-off study: tomography dimensioning, AO4ELT5 Proceedings <https://www.eso.org/~soberti/>
- Oberti, S., Kolb, J., Madec, P.-Y., et al. 2018, in *Adaptive Optics Systems VI*, eds. L. M. Close, L. Schreiber, & D. Schmidt, 10703, International Society for Optics and Photonics (SPIE), 469
- Oberti, S., Verinaud, C., Le Louarn, M., et al. 2019, in *6th International Conference on Adaptive Optics for Extremely Large Telescopes, AO4ELT 2019*, Québec, Canada
- Oppenheim, A. V., & Willsky, A. S. 1997, *Signals & Systems*, 2nd edn. (Prentice-Hall, Inc.)
- Paley, R. E., & Wiener, N. 1934, *Am. Math. Soc.*, **19**, 147
- Park, S. C., Park, M. K., & Kang, M. G. 2003, *IEEE Signal Process. Mag.*, **20**, 21
- Rigaut, F., & Gendron, E. 1992, *A&A*, **261**, 677
- Rigaut, F. J., Véran, J.-P., & Lai, O. 1998, *SPIE Conf. Ser.*, **3353**, 1038
- Rigaut, F., McDermid, R., Cresci, G., et al. 2020, *SPIE Conf. Ser.*, **11447**, 114471R
- Roddier, F. 1999, *Adaptive Optics in Astronomy* (New York: Cambridge University Press)
- Tsai, R. Y., & Huang, T. S. 1984, *Adv. Comput. Vis. Image Process.*, **1**, 317
- Verinaud, C. 2004, *Opt. Comm.*, **233**, 27
- Wang, L., Andersen, D., & Ellerbroek, B. 2012, *Appl. Opt.*, **51**, 3692
- Wang, Z., Chen, J., & Hoi, S. C. H. 2021, *IEEE Trans. Pattern Anal. Mach. Intell.*, **43**, 3365
- Wiener, N. 1930, *Acta Mathematica*, **55**, 117
- Wizinowich, P., Chin, J., Correia, C., et al. 2020, *SPIE Conf. Ser.*, **11448**, 114480E
- Willez, J., Abad, J. A., Abuter, R., et al. 2019, *A&A*, **629**, A41



## Appendix A: Signal reconstruction and optimal sampling grid positions

For the sake of completeness, we provide in this appendix mathematical arguments that help us estimate the number of sampling grids required for a reconstruction without aliasing, as a function of their relative spatial offsets. This study can be done with simple mathematical tools, if we consider a specific context that we describe below.

We propose to reconstruct a band-limited function  $f : \mathbb{R} \rightarrow \mathbb{R}$  sampled by a set of  $N$  grids with an interpolation formula of the form:

$$f(x) = \sum_{\ell=1}^N \sum_{n=-\infty}^{+\infty} f(u_n^\ell) s_\ell(x - u_n^\ell), \quad (\text{A.1})$$

where for each  $\ell \in \{1, \dots, N\}$ , the sequence  $u_n^\ell = \{u_n^\ell\}_{n \in \mathbb{Z}}$  gives the points where the function,  $f$ , is sampled. So, we have to find a suitable set of reconstruction functions  $s_1, \dots, s_N$ . The existence of such functions depends on the spectral properties of  $f$  and on characteristics of the set of sampling grids, such as their number, their step and their positions. One restriction lies in the fact that formula (A.1) is a convolution. It is a classical hypothesis, but it reduces the set of the possible reconstruction functions,  $s_1, \dots, s_N$ , and therefore the admissible positions of the grids. However, we can see that formula (A.1) leads to sinc-type reconstruction functions, which as such can be implemented at low numerical cost.

In the context of the PNS framework, we then suppose that all the grids have the same step  $h$ . However, we are free to set them at any position. Thus, they take the following form:

$$u_n^\ell = nh + u_0^\ell \quad \forall n \in \mathbb{Z} \quad \ell \in \{1, \dots, N\}, \quad (\text{A.2})$$

where  $u_0^\ell$  indicates their initial position, which is supposed to be different for each of them.

In the following, we establish the conditions, relating the maximal frequency of  $f$  to the characteristics of the sampling sequences, which ensure the existence of the functions  $s_1, \dots, s_N$  allowing us to reconstruct  $f$  with the formula (A.1). For a fixed maximal frequency, there are many possible options for the positions of the sampling grids, provided that we own a sufficient quantity of them. However, we are interested in finding the positions that need the smallest number of sampling grids as possible. As we go on to see, there is no surprise in this regard. With a set of uniform sampling grids, the meta-uniform sampling (Eq. (4)) is optimal: it allows us to reach a given maximal frequency with the lowest number of sampling grids.

### A.1. Reconstruction conditions

To find the reconstruction functions  $s_1, \dots, s_N$ , we solve Eq. (A.1) in the Fourier space. In this context, we assume the hypothesis that all involved functions admit a Fourier transform. So, we are looking for conditions that ensure the existence of some functions  $s_1, \dots, s_N$  such that:

$$\hat{f} = \hat{g} \quad \text{where} \quad g(x) := \sum_{\ell=1}^N \sum_{n=-\infty}^{+\infty} f(u_n^\ell) s_\ell(x - u_n^\ell) \quad \forall x \in \mathbb{R}.$$

For the sampling grids that we consider, the Fourier transform of  $g$  is given by:

$$\hat{g}(k) = C_0(k) \hat{f}(k) + \sum_{n \neq 0, n=-\infty}^{+\infty} C_n(k) \hat{f}(k - \nu_h n), \quad (\text{A.3})$$

where

$$C_n(k) := \frac{1}{h} \sum_{\ell=1}^N \hat{s}_\ell(k) e^{-2i\pi \nu_h n u_0^\ell} \quad \text{and} \quad \nu_h := \frac{1}{h}. \quad (\text{A.4})$$

This computation is not direct, but classical, and it can be found, for instance, in [Kohlenberg \(1953\)](#).

In what follows, we suppose that  $f$  is band-limited, hence  $\hat{f}(k)$  vanishes outside some symmetric interval  $K$ . That is, there exists a frequency,  $k_{\max} > 0$ , such that  $\hat{f}(k) = 0$  for all  $k \notin K$ , where  $K := (-k_{\max}, k_{\max})$ . In most practical situations  $\hat{f}$  does not necessarily have a bounded support. But,  $\hat{f}(k)$  may decrease rapidly as  $|k|$  goes to infinity, as does, for instance, the power spectrum density of turbulent wavefronts following a Kolmogorov power law.

The translated copies  $\hat{f}(k - \nu_h n)$  of  $\hat{f}$ , that appear in Eq. (A.3), may have some non zero values in the set  $K$  overlapping  $\hat{f}$  in this set, and hence producing aliasing. This happens if for some  $n \neq 0$  a set  $K_n := (-k_{\max} + \nu_h n, k_{\max} + \nu_h n)$  satisfies  $K \cap K_n \neq \emptyset$ . It is the case if and only if  $|n| < n_0$ , where

$$n_0 := \left\lceil \frac{2k_{\max}}{\nu_h} \right\rceil = \left\lceil \frac{k_{\max}}{\nu} \right\rceil, \quad (\text{A.5})$$

and  $\lceil \cdot \rceil$  is the ceiling function and  $\nu = 1/2h$  is the native Nyquist frequency. This can be reformulated in the following way: if for some integer  $n_0$  the frequency  $k_{\max}$  satisfies  $(n_0 - 1)\nu < k_{\max} \leq n_0\nu$ , then  $K \cap K_n \neq \emptyset$ , if and only if  $|n| < n_0$ .

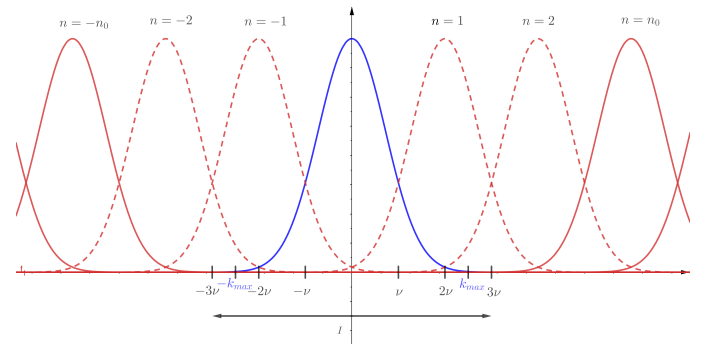
It follows, that in  $K$ , but also in the bigger set:

$$I := [-n_0\nu, n_0\nu] \supset K,$$

the formula (A.3) is expressed as:

$$\hat{g}(k) = C_0(k) \hat{f}(k) + \sum_{n \neq 0, n=-n_0+1}^{n_0-1} C_n(k) \hat{f}(k - \nu_h n) \quad \forall k \in I, \quad (\text{A.6})$$

since  $\hat{f}(k - \nu_h n) = 0$  for every  $k \in I$  if  $|n| \geq n_0$ , see Figure A.1.



**Fig. A.1.** Fourier transform  $\hat{f}$  (blue) and its translated copies (red). Here,  $2\nu < k_{\max} \leq 3\nu$  and therefore  $n_0 = 3$ . The translated copies  $\hat{f}(k - \nu_h n)$  overlap  $\hat{f}$  for  $n \in \{-2, -1, 1, 2\}$  (dashed curves). On the interval  $I = [-3\nu, 3\nu]$ , there is no contribution of  $\hat{f}(k - \nu_h n)$  for any  $|n| \geq n_0$ .

To ensure that  $\hat{g} = \hat{f}$  in  $I$ , and therefore in  $K$ , we can impose for every  $k \in I$  that  $C_0(k) = 1$  and  $C_n(k) = 0$  for all



$n \in \{-n_0 + 1, \dots, n_0 - 1\} \setminus \{0\}$ , that is,

$$\sum_{\ell=1}^N \hat{s}_\ell(k) = h \quad \forall k \in I, \quad (\text{A.7})$$

$$\sum_{\ell=1}^N \hat{s}_\ell(k) e^{2i\pi n_\ell u_0^\ell} = 0 \quad \forall n \in \{-n_0 + 1, \dots, n_0 - 1\} \setminus \{0\} \quad \forall k \in I. \quad (\text{A.8})$$

Now, for  $k \notin I$ , we have  $\hat{f}(k) = 0$  and therefore we want  $\hat{g}(k) = 0$ . To cancel the contributions of all the translated copies of  $\hat{f}$  outside of  $K$ , we can simply set:

$$\hat{s}_1(k) = \dots = \hat{s}_N(k) = 0 \quad \forall k \notin I. \quad (\text{A.9})$$

**Remark:** Eq. (A.7) and Eq. (A.8) are sufficient conditions to obtain  $\hat{g} = \hat{f}$  in the set  $I$  and to reconstruct  $f$  using formula (A.1). However, it may be possible to propose weaker but more elaborated conditions. For the sake of simplicity, we can restrict the study to Eqs. (A.7) and (A.8).

We can sum up our results and introduce more suitable notations for the forthcoming study of the system given by Eq. (A.7) and Eq. (A.8). We consider a band-limited function  $f$ , whose Fourier transform vanishes outside of  $(-k_{\max}, k_{\max})$ , and  $N$  sampling sequences of step  $h$  and initial positions  $u_0^1, \dots, u_0^N$  (see Eq. (A.2)). If we introduce the quantities:

$$\beta_\ell := e^{2i\pi n_\ell u_0^\ell} = e^{2i\pi u_0^\ell/h} \quad \forall \ell \in \{1, \dots, N\},$$

then the system defined by Eq. (A.7) and Eq. (A.8) is of the form  $AX = Y$ , where  $X$  is a vector of size  $N$ ,  $A$  is the  $(2n_0 - 1) \times N$  matrix defined by:

$$A := \begin{pmatrix} \beta_1^{-n_0+1} & \beta_2^{-n_0+1} & \dots & \beta_N^{-n_0+1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^{n_0-1} & \beta_2^{n_0-1} & \dots & \beta_N^{n_0-1} \end{pmatrix} \quad \text{and} \quad Y := \begin{pmatrix} 0 \\ \vdots \\ h \\ \vdots \\ 0 \end{pmatrix}. \quad (\text{A.10})$$

If  $X = (x_1, \dots, x_N)$  is a solution of the system A.10, and  $\hat{s}_1, \dots, \hat{s}_N$  are defined for any  $\ell \in \{1, \dots, N\}$  by:

$$\hat{s}_\ell(k) = x_\ell \quad \forall k \in I \quad \text{and} \quad \hat{s}_\ell(k) = 0 \quad \forall k \notin I, \quad (\text{A.11})$$

then  $\hat{g} = \hat{f}$ . It follows that Eq. (A.1) holds, with the reconstruction functions:

$$s_\ell(x) = \frac{x_\ell n_0}{h} \text{sinc}\left(\frac{\pi n_0 x}{h}\right) \quad \forall \ell \in \{1, \dots, N\}. \quad (\text{A.12})$$

We can offer a simple example and confirm the Nyquist–Shannon Theorem anew at the same time.

**Example:** By the Nyquist–Shannon Theorem, we know that, if  $k_{\max} \leq 1/2h = \nu$ , a sampling sequence of step,  $h$ , allows us to perfectly reconstruct  $f$ . We can re-prove that result. Indeed, if  $k_{\max} \leq 1/2h$ , then formula (A.5) gives  $n_0 = 1$  and the system (A.10) to solve is expressed as:

$$x_1 + \dots + x_N = h.$$

If  $N \geq 2$ , we can produce an infinite number of solutions of this equation, and we have as many possible choices for  $\hat{s}_1, \dots, \hat{s}_N$ .

However, a single sampling grid ( $N = 1$ ), is enough to ensure the existence of a solution, which is  $x_1 = h$ . So, we let

$$\hat{s}_1(k) = h \quad \forall k \in I \quad \text{and} \quad \hat{s}_1(k) = 0 \quad \forall k \notin I,$$

where  $I = [-1/2h, 1/2h]$ . By applying Eq. (A.12) and adopting Eq. (A.1), we obtain

$$s_1(x) = \text{sinc}(\pi x/h) \quad \text{and}$$

$$f(x) = \sum_{n=-\infty}^{+\infty} f(nh + u_0^1) \text{sinc}\left(\frac{\pi}{h}(x - nh - u_0^1)\right),$$

as expected.

In the following, we study the conditions on  $N$  and  $u_0^1, \dots, u_0^N$  that ensure the existence of a solution to the system (A.10). As explained, if such a system has a solution, then we can reconstruct  $f$  using formula (A.1). If the system does not have solution, then there is no functions  $\hat{s}_1, \dots, \hat{s}_N$ , that satisfies Eq. (A.7) and Eq. (A.8). This does not necessarily mean that  $f$  cannot be reconstructed with formula (A.1), since Eq. (A.7) and Eq. (A.8) are sufficient, but not necessary, conditions to obtain  $\hat{g} = \hat{f}$ .

## A.2. Optimality of the uniform sampling

The system (A.10) has  $2n_0 - 1$  equations with  $N$  unknowns. The number of equations is imposed by the highest frequency  $k_{\max}$  of the signal  $f$  and by the step  $h$  of the sampling grids, through the formula (A.5). The number of unknowns is equal to the number  $N$  of sampling grids that we own. The coefficients of the system depend on  $h$  and can be tuned by our choice of the initial positions  $u_0^1, \dots, u_0^N$  of the sampling grids. Here, we are interested in the initial positions which ensure the existence of a solution to Eq. (A.10), with the smallest number of sampling grids  $N$  as possible.

### A.2.1. Minimal number of sampling grids

If  $N = 2n_0 - 1$ , then  $A$  is a square matrix with a Vandermonde determinant. This determinant vanishes if and only if  $\beta_\ell = \beta_{\ell'}$  for some  $\ell \neq \ell'$ . This occurs if and only if  $u_0^\ell = u_0^{\ell'} + qh$  for some  $q \in \mathbb{Z}$ , that is, if the two grids  $\ell$  and  $\ell'$  have the same points (see Eq. A.2). Therefore, if the sampling grids are all different, then the matrix  $A$  is invertible. In such a case, the system (A.10) has a solution that is unique for each possible position of  $u_0^1, \dots, u_0^N$ . We can then reconstruct our signal choosing the functions  $\hat{s}_1, \dots, \hat{s}_N$  according to Eq. (A.11). We conclude that with  $N = 2n_0 - 1$  grids, we can reconstruct a signal such that  $(n_0 - 1)\nu < k_{\max} \leq n_0\nu$ , even if the initial positions of the grids are chosen randomly.

However, for aptly chosen positions of  $u_0^1, \dots, u_0^N$ , we can reduce the number of independent equations of the system and obtain a solution even if  $N < 2n_0 - 1$ . Indeed, we may suppose that we have  $N = n_0$  sampling grids of the form:

$$u_n^\ell = nh + u_0^1 + (\ell - 1)\frac{h}{N} \quad \forall \ell \in \{1, \dots, N\}. \quad (\text{A.13})$$

Then, for every  $\ell \in \{1, \dots, N\}$  and  $j \in \{-n_0 + 1, \dots, -1\}$ , we have

$$\beta_\ell^{j+n_0} = e^{2i\pi(j+N)u_0^\ell/h} = \beta_\ell^j e^{2i\pi N u_0^1/h} e^{2i\pi(j-1)h/N} = \beta_1^N \beta_\ell^j.$$

This implies that to multiply the  $n_0 - 1$  first lines of  $AX = Y$  by  $\beta_1^N$  gives a system where each of the  $n_0 - 1$  first equations is equal

to one of the  $n_0 - 1$  last equations. Therefore, if  $X = (x_1, \dots, x_N)$  is a solution for:

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \beta_1 & \beta_2 & \cdots & \beta_N \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^{n_0-1} & \beta_2^{n_0-1} & \cdots & \beta_N^{n_0-1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} h \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad (\text{A.14})$$

then  $X$  is also a solution of  $AX = Y$ . The square system (A.14) has a Vandermonde determinant. Since the points of the grids (A.13) are all different, the system (A.14) is invertible. We can verify that its unique solution is given by  $x_1 = x_2 = \cdots = x_N = h/N$ . So, we re-prove that with  $N = n_0$  grids chosen according to Eq. (A.13), we can reconstruct a signal such that  $(n_0 - 1)\nu < k_{\max} \leq n_0\nu$ .

Now, we may wonder if there is some smarter way to select the positions of the sampling grids, which would allow us to further reduce their number. The answer is negative. Indeed, if  $N \leq n_0 - 1$ , we can extract, from Eq. (A.10), the subsystem:

$$\begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_N \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^N & \beta_2^N & \cdots & \beta_N^N \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \quad (\text{A.15})$$

which must be satisfied by any solution of Eq. (A.10). As Eq. (A.15) is an invertible homogeneous system, its unique solution is  $X = (0, 0, \dots, 0)$ , which is not a solution of Eq. (A.10). Thus, Eq. (A.10) has no solution if  $N \leq n_0 - 1$ .

### A.2.2. Unicity

We have shown that  $N = n_0$  is the minimal number of sampling grids for the existence of a solution to the system (A.10). In particular, if we choose our grids according to Eq. (A.13), then (A.10) has a solution. At this point, we can verify that for any other choice of the position of  $n_0$  grids, the system (A.10) has no solution.

To that aim, we look for a necessary condition for the rank of  $A$  to be equal to the rank of  $A|Y$ , where  $A|Y$  is the matrix  $A$  with an extra row equal to  $Y$ . First, we have  $\text{rank}(A) = N$ , since, on the one hand, the number of rows  $N = n_0$  of  $A$  is smaller than its number of lines  $2n_0 - 1$ ; and, on the other hand,  $A$  contains a  $N \times N$  Vandermonde invertible matrix. Thus, it is enough to look for a necessary condition for  $\text{rank}(A|Y) = N$ .

As  $N = n_0$ ,  $A|Y$  is a  $(2N - 1) \times (N + 1)$  matrix, from which we can extract a collection of  $N - 1$  matrices  $B_1, \dots, B_{N-1}$  of size  $(N + 1) \times (N + 1)$  and defined by:

$$B_i := \begin{pmatrix} \beta_1^{-i} & \beta_2^{-i} & \cdots & \beta_N^{-i} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 & h \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta_1^{N-i} & \beta_2^{N-i} & \cdots & \beta_N^{N-i} & 0 \end{pmatrix} \quad \forall i \in \{1, \dots, N - 1\}.$$

We can have  $\text{rank}(A|Y) = N$ , only if the  $\det(B_i) = 0$  for all  $i \in \{1, \dots, N - 1\}$ . A Laplace expansion along the last row gives:

$$\det(B_i) = (-1)^{N+i+2} h \det(C_i) \prod_{\ell=1}^N \beta_\ell^{-i} \quad \text{where}$$

$$C_i := \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^{i-1} & \beta_2^{i-1} & \cdots & \beta_N^{i-1} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^N & \beta_2^N & \cdots & \beta_N^N \end{pmatrix} \quad \forall i \in \{1, \dots, N - 1\}.$$

Let  $P(X) = a_N X^N + \cdots + a_i X^i + \cdots + a_0$  be the polynomial defined by

$$P(X) := \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta_1^i & \beta_2^i & \cdots & \beta_N^i & X^i \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta_1^N & \beta_2^N & \cdots & \beta_N^N & X^N \end{vmatrix}.$$

Then, a Laplace expansion along the last row gives  $a_i = (-1)^{i+N+1} \det(C_i)$  for all  $i \in \{1, \dots, N - 1\}$ . Therefore, we have  $\det(B_i) = 0$  for all  $i \in \{1, \dots, N - 1\}$  if and only if

$$P(X) = a_N X^N + a_0.$$

The coefficients  $a_N$  and  $a_0$  are obtained by Laplace expansion, and if we denote  $V_N(\beta_1, \dots, \beta_N)$  the Vandermonde determinant of size  $N$ , then:

$$P(X) = V_N(\beta_1, \dots, \beta_N) \left( X^N + (-1)^{N+2} \prod_{\ell=1}^N \beta_\ell \right).$$

On the other hand, we have  $P(\beta_\ell) = 0$  for all  $\ell \in \{1, \dots, N\}$ , which gives:

$$\beta_1^N = \beta_2^N = \cdots = \beta_N^N = (-1)^{N+1} \prod_{\ell=1}^N \beta_\ell.$$

It follows that  $(\beta_\ell / \beta_1)^N = 1$  for all  $\ell \in \{1, \dots, N\}$ . In other words, each  $\beta_\ell / \beta_1$  is a  $N^{\text{th}}$  root of the unity. Since all the  $\beta_\ell$  are supposed different, we have

$$\beta_\ell = \beta_1 e^{2i\pi(\ell-1)/N} = e^{2i\pi(u_0^1/h + (\ell-1)/N)} \quad \forall \ell \in \{1, \dots, N\}.$$

Recalling that  $\beta_\ell = e^{2i\pi u_0^\ell/h}$ , we obtain

$$u_0^\ell = u_0^1 + (\ell - 1) \frac{h}{N} \mod h \quad \forall \ell \in \{1, \dots, N\}.$$

We conclude that, when  $N = n_0$ , the sampling grids must satisfy Eq. (A.13) to ensure the existence of a solution of the system (A.10).

### A.2.3. Conclusion

To sum up our findings, with  $N = 2n_0 - 1 = 2 \lceil k_{\max}/\nu \rceil - 1$  sampling grids of step  $h$ , we can always reconstruct a signal,  $f$ , with a maximal frequency of  $k_{\max}$ . The sampling grid positions can be chosen at random, provided that the grids are all different. However, the same performance can be reached with only  $N = n_0$  sampling grids, if their union is a uniform sampling grid of step  $h/N$ , that is, if they satisfy Eq. (A.13) modulo a translation of

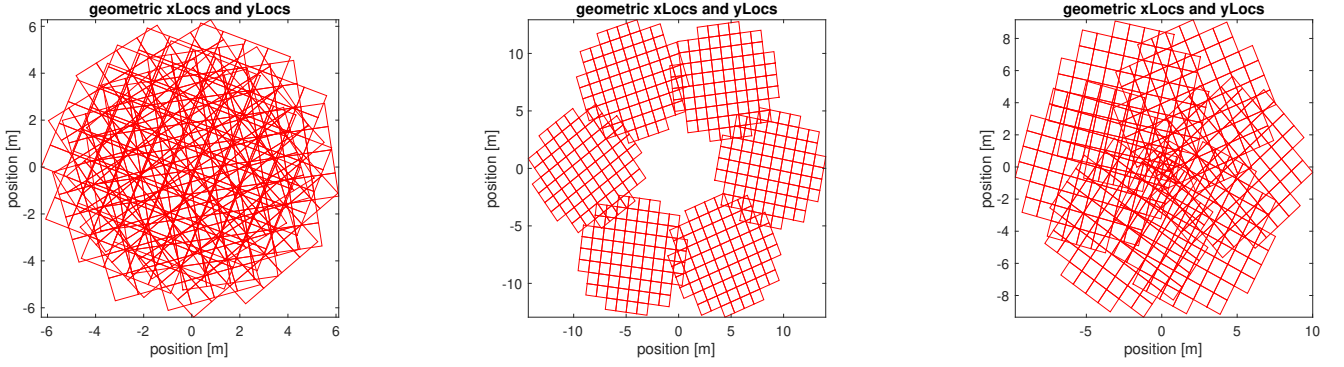
a size  $h$ . On the other hand, if we have fewer than  $n_0$  sampling grids, or  $n_0$  sampling grids that do not satisfy Eq. (A.13), then there is no set of reconstruction functions  $s_1, \dots, s_N$  which satisfy Eq. (A.7) and Eq. (A.8). Since these conditions might be too strong, we cannot draw a firm conclusion that the signal cannot be reconstructed, since it is possible that more elaborate approaches and reconstruction functions could offer a more promising result. Our study suggests nevertheless that the meta-uniform sampling Eq. (A.13) is optimal, because it allows us to reconstruct the signal without aliasing with the smallest number of sampling grids as possible. It is therefore the safest way to make use of a limited number of sampling grids.

## Appendix B: Non-uniform sampling schemes

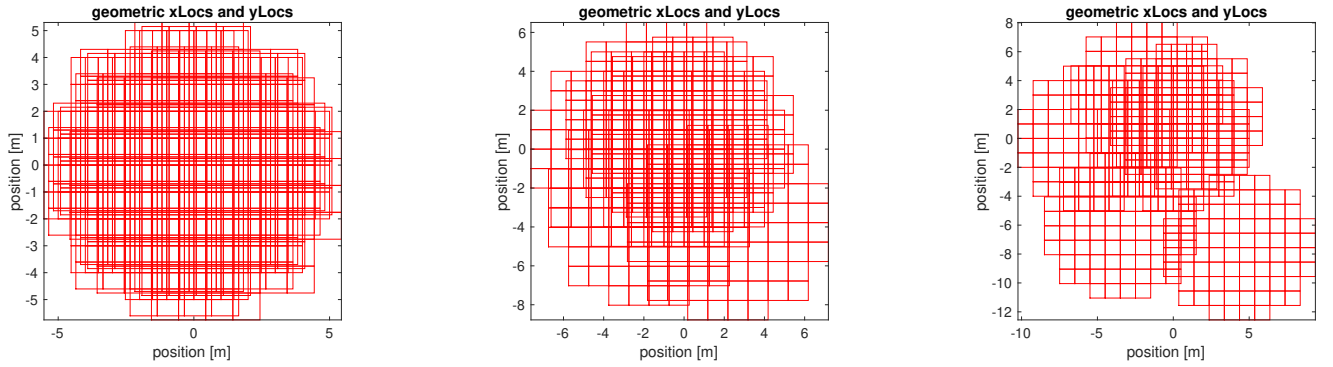
We have shown that both uniform and non-uniform sampling schemes allow for the reconstruction of a band-limited signal without aliasing, under some conditions presented in this paper. In addition to offsets and rotations, the AO design offers several other sampling schemes that facilitate SR. A non-exhaustive list of recommendations is as follows:

- differently rotate and offset (possibly stretch) the WFS sampling arrays by non-integer multiples of each other (Figure B.1);
- use asymmetric guide star asterisms projected on-sky (Figure B.2);
- use WFSs with different numbers and sizes of sub-apertures (Figure B.3);
- allowing the sub-aperture sizes in each lenslet array to vary.

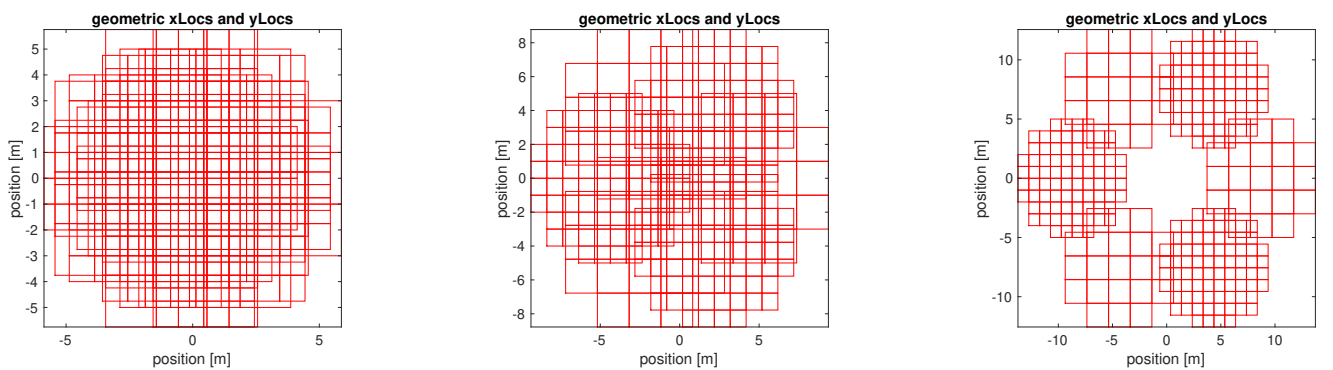
In either case, the unique phase information collected by each WFS can be used to separate the aliased high-frequency from the low-frequency content of interest and the higher-resolution wavefront can be accurately reconstructed.



**Fig. B.1.** Six differentially rotated WFS configurations shown at three different arbitrary conjugation altitudes creating growing separations (highly exaggerated for illustration only).



**Fig. B.2.** Six regular WFS configurations with a spiral shaped asterism shown at three different conjugation altitudes creating growing separations.



**Fig. B.3.** Six WFS configurations with two kinds of sub-aperture, one being the double in size of the other, shown at three different conjugation altitudes creating growing separations.